AN A POSTERIORI ERROR ESTIMATOR FOR DISCONTINUOUS GALERKIN METHODS
FOR NON-STATIONARY CONVECTION-DIFFUSION PROBLEMS

ANDREA CANGIANI, EMMANUIL H. GEORGOULIS AND STEPHEN METCALFE

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF LEICESTER, UNIVERSITY
ROAD, LEICESTER LE1 7RH, UNITED KINGDOM

Abstract. This work is concerned with the derivation of a robust a posteriori error estimator for a discontinuous Galerkin method discretisation of linear non-stationary convection-diffusion initial/boundary value problems and with the implementation of a corresponding adaptive algorithm. More specifically, we derive a posteriori bounds for the error in the $L^2(H^1)$-type norm for an interior penalty discontinuous Galerkin (dG) discretisation in space and a backward Euler discretisation in time. An important feature of the estimator is robustness with respect to the Péclet number of the problem which is verified in practice by a series of numerical experiments. Finally, an adaptive algorithm is proposed utilising the error estimator. Optimal rate of convergence of the adaptive algorithm is observed in a number of test problems. discontinuous Galerkin methods; a posteriori error estimators; time-dependent convection-diffusion equation.

1. Introduction. The interaction between convection and diffusion modelled by initial/boundary value problems involving partial differential equations (PDEs) poses a number of challenges in the context of their numerical approximation. Indeed, stationary convection-dominated convection-diffusion type problems admit analytical solutions of a multiscale nature that can contain steep gradients, usually termed boundary or interior layers, depending upon their location in the computational domain. The accurate and efficient numerical resolution of such steep layers is a challenge as their exact location cannot be, in general, known a priori. In the special cases where the location of boundary or interior layers is known, structured grids have been successfully employed ([20]). Ultimately, in non-stationary convection-diffusion equations, the nature of the solution (including layers) may vary throughout the domain as time progresses. This renders the use of adaptive algorithms an attractive proposition for the accurate and efficient numerical approximation of this class of problems. As adaptive algorithms are usually based on suitable a posteriori error estimators, the formulation of estimators that can robustly estimate both the temporal and spatial nature of the error are of particular interest.

A posteriori error estimation for stationary linear equations is now relatively well understood; for pure diffusion problems there is a huge array of estimators available for a wide variety of different types of finite element discretisations ([30, 1]) and for dG methods in particular ([16, 6, 14]). For stationary convection-diffusion equations, the quest for robust a posteriori error estimators (in the sense that they are independent of the Péclet number of the problem) has seen recent advancements in various contexts ([27, 29, 17, 21, 23, 24]).

A posteriori error estimators for non-stationary linear convection-diffusion equations are also available for various discretisations ([15, 7, 28, 10, 2, 3, 26, 12, 18]). A posteriori error estimators for (spatial) dG methods for non-stationary pure diffusion problems can be found in [13, 11, 25].

This work is concerned with the derivation and implementation of an a posteriori estimator in the $L^2(H^1)$-norm for the backward Euler discretisation in time (zero-th order dG method) and the interior penalty dG discretisation in space of the non-stationary linear convection-diffusion equation with variable coefficients. The derivation of the a posteriori bound utilises the elliptic reconstruction technique ([19, 13])

arXiv:1211.3632v1 [math.NA] 15 Nov 2012

which allows the use of the robust elliptic error estimator from [23]. The temporal residual is also treated so as to ensure that robustness with respect to the Péclet number of the problem is maintained. The resulting a posteriori error estimator is shown numerically to be reliable in both time and space. The a posteriori estimator derived below can be viewed as the analogue of the one presented in [28] when the spatial discretisation is a dG method.

The a posteriori estimator is used to drive an adaptive algorithm which is numerically assessed in two ways. First, the savings with respect to the degrees of freedom by using space-time adaptivity are shown in a series of test problems. Second, the rates of convergence of the space-time adaptive algorithm are computed. The adaptive algorithm appears to convergence optimally in both space and time when the spatial polynomial degree, p, satisfies $p \geq 2$. Somewhat surprisingly, for $p = 1$, the a posteriori estimator appears to converge at a suboptimal rate when the mesh changes between time steps. The nature of this degeneracy for $p = 1$ remains unclear to the authors: it is in fact possible that the method itself is suboptimal in time when mesh change is allowed. It is also unclear if this is a special phenomenon of the discretisation presented in this work or a more general phenomenon extending to other discretisations of parabolic problems. This phenomenon will be investigated further elsewhere.

The remainder of this work is structured as follows. The function space setting and the model problem are given in Section 2. Section 3 describes the discretisation of the problem. In Section 4 we state an a posteriori error bound for the stationary elliptic problem. Sections 5 and 6 contain the derivation of the a posteriori bounds for the non-stationary convection-diffusion model problem in the semi-discrete and the fully-discrete settings, respectively. Section 7 describes a space-time adaptive algorithm driven by the a posteriori estimators derived in the previous section. A series of numerical experiments are presented in Section 8 and the final conclusions are drawn in Section 9.

2. Model problem. Let the computational domain $\Omega \subset \mathbb{R}^2$ be a bounded Lipschitz polygon with boundary $\partial\Omega$.

We denote the standard L^2-inner product on Ω by $(\cdot, \cdot)$ and the standard L^2-norm on Ω by $|| \cdot ||$. For $1 \leq p \leq +\infty$, we define the spaces $L^p(0, T; X)$ (where X is a real Banach space with norm $\| \cdot \|_X$) that consist of all measurable functions $v : [0, T] \to X$ for which:

$$
\begin{aligned}
&\|v\|_{L^p(0,T;X)} := \Big(\int_0^T \|v(t)\|_X^p dt \Big)^{1/p} < +\infty, \quad \text{for} \quad 1 \leq p < +\infty, \\
&\|v\|_{L^\infty(0,T;X)} := \operatorname*{ess\,sup}_{0 \leq t \leq T} \|v(t)\|_X < +\infty, \quad \text{for} \quad p = +\infty.
\end{aligned}
\tag{2.1}
$$

We also define $H^1(0, T, X) := \{u \in L^2(0, T; X) : u_t \in L^2(0, T; X)\}$. Finally, we denote by $C(0, T; X)$ and $C^{0,1}(0, T; X)$, respectively, the spaces of continuous and Lipschitz-continuous functions $v : [0, T] \to X$ such that:

$$
\begin{aligned}
&||v||_{C(0,T;X)} = \max_{0 \leq t \leq T} ||v(t)||_X < \infty, \\
&||v||_{C^{0,1}(0,T;X)} = \max \Big\{ ||v||_{C(0,T;X)}, ||\frac{\partial v}{\partial t}||_{C(0,T,X)} \Big\} < \infty.
\end{aligned}
\tag{2.2}
$$

2

We consider the model problem of finding $u : \Omega \to \mathbb{R}$ such that:

$$\frac{\partial u}{\partial t} - \varepsilon \Delta u + \mathbf{a} \cdot \nabla u + bu = f \qquad \text{in } \Omega \times (0, T],$$
$$u = 0 \qquad \text{on } \partial\Omega \times (0, T], \qquad (2.3)$$
$$u(\cdot, 0) = u_0 \qquad \text{in } \bar{\Omega}.$$

We make the following assumptions: $u_0 \in L^2(\Omega)$, $f \in C(0, T; L^2(\Omega))$, $\varepsilon \in (0, 1]$, $\mathbf{a} \in C(0, T; W^{1,\infty}(\Omega))^2$, and $b \in C(0, T; L^\infty(\Omega))$.

It is assumed that the length of $\mathbf{a} \equiv \mathbf{a}(\mathbf{x}, \mathbf{t})$ and the area of Ω are of order one, possibly up to rescaling, so that ϵ^{-1} can be taken to be the Péclet number of the problem. For simplicity, we assume that there are constants $\beta \geq 0$ and $c_* \geq 0$ such that:

$$b - \frac{1}{2} \nabla \cdot \mathbf{a} \geq \beta \quad \text{a.e. in } \Omega \times [0, T], \qquad || - \nabla \cdot \mathbf{a} + b||_{C(0,T;L^\infty(\Omega))} \leq c_* \beta. \qquad (2.4)$$

The weak form of (2.3) then reads: find $u \in L^2(0, T; H_0^1(\Omega)) \cap H^1(0, T; L^2(\Omega))$ such that for each $t \in (0, T]$ we have

$$\int_\Omega \frac{\partial u}{\partial t} v \, dx + \int_\Omega (\varepsilon \nabla u \cdot \nabla v + \mathbf{a} \cdot \nabla u v + buv) \, dx = \int_\Omega fv \, dx \quad \forall v \in H_0^1(\Omega). \qquad (2.5)$$

Under the regularity assumptions above, we have that $u \in C(0, T; L^2(\Omega))$.

3. Discontinuous Galerkin method. Let the mesh $\zeta = \{K\}$ be a shape-regular subdivision of Ω, with K denoting a generic element. We assume that the subdivision ζ is constructed via affine mappings $F_K : \hat{K} \to K$ with non-singular Jacobian where $\hat{K}$ is the reference triangle or the reference square. The mesh is allowed to contain a uniformly fixed number of regular hanging nodes per edge. We define the finite element space

$$V_h \equiv V_h(\zeta) := \{v \in L^2(\Omega) : v|_K \circ F_K \in \mathcal{P}^p(\hat{K}), K \in \zeta\}, \qquad (3.1)$$

where $\mathcal{P}^p(K)$ is the space of polynomials of total degree p if $\hat{K}$ is the reference triangle, or the space of polynomials of degree p in each variable if $\hat{K}$ is the reference square.

Denote by $\mathcal{E}(\zeta)$ the set of all edges in the triangulation ζ and $\mathcal{E}^{int}(\zeta) \subset \mathcal{E}(\zeta)$ the subset of all interior edges. We also denote the diameter of an element K by h_K and the length of an edge E by h_E. The outward unit normal on the boundary ∂K of an element K is denoted by $\mathbf{n}_K$. Given an edge $E \in \mathcal{E}^{int}(\zeta)$ shared by two elements K and K', a vector field $\mathbf{v} \in [H^{1/2}(\Omega)]^2$ and a scalar field $v \in H^{1/2}(\Omega)$, we define jumps and averages of $\mathbf{v}$ and v across E by:

$$\{\mathbf{v}\} = \frac{1}{2}(\mathbf{v}|_{\bar{K}} + \mathbf{v}|_{\bar{K}'}), \qquad [\mathbf{v}] = \mathbf{v}|_{\bar{K}} \cdot \mathbf{n}_K + \mathbf{v}|_{\bar{K}'} \cdot \mathbf{n}_{K'},$$
$$\{v\} = \frac{1}{2}(v|_{\bar{K}} + v|_{\bar{K}'}), \qquad [v] = v|_{\bar{K}} \mathbf{n}_K + v|_{\bar{K}'} \mathbf{n}_{K'}. \qquad (3.2)$$

If $E \subset \partial\Omega$, we set $\{\mathbf{v}\} = \mathbf{v}$, $[\mathbf{v}] = \mathbf{v} \cdot \mathbf{n}$, $\{v\} = v$ and $[v] = v\mathbf{n}$, with $\mathbf{n}$ denoting the outward unit normal to the boundary $\partial\Omega$.

We define the inflow and outflow parts of the boundary $\partial\Omega$ at the time t, respectively, by:

$$\partial\Omega_{in}^t = \{x \in \partial\Omega \mid \mathbf{a}(x, t) \cdot \mathbf{n}(x) < 0\}, \qquad \partial\Omega_{out}^t = \{x \in \partial\Omega \mid \mathbf{a}(x, t) \cdot \mathbf{n}(x) \geq 0\}. \quad (3.3)$$

The inflow and outflow parts of an element K at time t are similarly defined as:

$$\partial K_{in}^t = \{x \in \partial K \mid \mathbf{a}(x,t) \cdot \mathbf{n}_K(x) < 0\}, \qquad \partial K_{out}^t = \{x \in \partial K \mid \mathbf{a}(x,t) \cdot \mathbf{n}_K(x) \geq 0\}. \tag{3.4}$$

The semi-discrete discontinuous Galerkin approximation to (2.5) then reads as follows. For $t = 0$, set $u_h(0) \in V_h$ to be a projection of u_0 onto V_h. For $t \in (0,T]$, find $u_h \in C^{0,1}(0,T;V_h)$ such that

$$(\frac{\partial u_h}{\partial t}, v_h) + B(t; u_h, v_h) + K_h(u_h, v_h) = (f, v_h) \qquad \forall v_h \in V_h, \tag{3.5}$$

with

$$B(t; w, v) \equiv B(\zeta; t; w, v)$$
$$:= \sum_{K \in \zeta} \int_K (\varepsilon \nabla w - \mathbf{a}w) \cdot \nabla v + (b - \nabla \cdot \mathbf{a})wv \, dx + \sum_{E \in \mathcal{E}(\zeta)} \frac{\varepsilon \gamma}{h_E} \int_E [w] \cdot [v] \, ds$$
$$+ \sum_{K \in \zeta} \left(\int_{\partial K_{out}^t \cap \partial \Omega_{out}^t} \mathbf{a} \cdot \mathbf{n}_K wv \, ds + \int_{\partial K_{out}^t \setminus \partial \Omega} \mathbf{a} \cdot \mathbf{n}_K w(v|_{\bar{K}} - v|_{\bar{K}'}) \, ds \right),$$
$$K_h(w, v) \equiv K_h(\zeta; w, v) := - \sum_{E \in \mathcal{E}(\zeta)} \int_E \{\varepsilon \nabla w\} \cdot [v] + \{\varepsilon \nabla v\} \cdot [w] \, ds,$$
$$\tag{3.6}$$

where $\bar{K} \cap \bar{K}' = E \subset \partial K_{out}^t \setminus \partial \Omega$. In standard fashion, the penalty parameter $\gamma > 0$ is chosen large enough so that the operator $B + K_h$ is coercive. Moreover, for simplicity of the presentation, we assume $\gamma > 1$ so that the constants in the subsequent discussion are independent of it.

We note that the bilinear form K_h is not well-defined for arguments in $H_0^1(\Omega)$, but the bilinear form B is and for $u, v \in H_0^1(\Omega)$ and $t \in (0,T]$, we have

$$B(t; u, v) = \int_\Omega (\varepsilon \nabla u \cdot \nabla v + \mathbf{a} \cdot \nabla uv + buv) \, dx. \tag{3.7}$$

In light of this, the weak problem (2.5) can be rewritten for each $t \in (0,T]$ as

$$(\frac{\partial u}{\partial t}, v) + B(t; u, v) = (f, v) \qquad \forall v \in H_0^1(\Omega). \tag{3.8}$$

We shall also consider a full discretisation of problem (2.5) by using a backward Euler method to approximate the time derivative.

To this end, consider a subdivision of $[0,T]$ into time intervals of lengths $\tau_1, \tau_2, ..., \tau_n$ such that $\sum_{j=1}^n \tau_j = T$ for some $n \geq 1$ and set $t^0 = 0$ and $t^k := \sum_{j=1}^k \tau_k$. Denote an initial triangulation by ζ^0. We associate to each time step $k > 0$ a triangulation ζ^k which is assumed to have been obtained from ζ^{k-1} by locally refining and coarsening ζ^{k-1}. This restriction upon mesh change is made to avoid degradation of the finite element solution, cf. [9]. To each mesh ζ^k, we assign the finite element space $V_h^k = V_h(\zeta^k)$ given by (3.1).

We set $B^k(t; \cdot, \cdot) := B(\zeta^k; t; \cdot, \cdot)$ and $K_h^k(\cdot, \cdot) := K_h(\zeta^k; \cdot, \cdot)$. Moreover, when the last two arguments of B are in $H_0^1(\Omega)$, we simply denote the bilinear form by B. We also denote $f(., t^k) = f^k$, $\mathbf{a}(., t^k) = \mathbf{a}^k$, and $b(., t^k) = b^k$ for brevity.

The fully-discrete dG method then reads as follows. Set u_h^0 to be a projection of u_0 onto V_h^0. For $k = 0,...,n-1$, find $u_h^{k+1} \in V_h^{k+1}$ such that

$$(\frac{u_h^{k+1} - u_h^k}{\tau_{k+1}}, v_h^{k+1}) + B^{k+1}(t^{k+1}; u_h^{k+1}, v_h^{k+1}) + K_h^{k+1}(u_h^{k+1}, v_h^{k+1}) = (f^{k+1}, v_h^{k+1})$$

$$\forall v_h^{k+1} \in V_h^{k+1}. \tag{3.9}$$

We shall take u_h^0 to be the orthogonal L^2-projection of u_0 onto V_h^0, although other projections onto V_h^0 can also be used.

4. Error bounds for the dG method for the stationary problem. To analyse the error, we introduce the following quantities:

$$|||u|||_\zeta := \left(\sum_{K \in \zeta} (\varepsilon ||\nabla u||_{L^2(K)}^2 + \beta ||u||_{L^2(K)}^2) + \sum_{E \in \mathcal{E}(\zeta)} \frac{\varepsilon \gamma}{h_E} ||[u]||_{L^2(E)}^2 \right)^{1/2},$$

$$|u|_{A,\zeta} := \left(\left(\sup_{v \in H_0^1(\Omega) \backslash \{0\}} \frac{\int_\Omega \mathbf{a} u \cdot \nabla v \, dx}{|||v|||} \right)^2 + \sum_{E \in \mathcal{E}(\zeta)} \left(\beta h_E + \frac{h_E}{\varepsilon} \right) ||[u]||_{L^2(E)}^2 \right)^{1/2}. \tag{4.1}$$

We note that $||| \cdot |||_\zeta$ and $| \cdot |_{A,\zeta}$ define norms on $H_0^1(\Omega) + V_h$. When no confusion is likely to occur we shall suppress the dependence on the mesh ζ and write $||| \cdot |||$ and $| \cdot |_A$.

Throughout this work, the symbols $\lesssim$ and $\gtrsim$ are used to denote inequalities true up to a positive constant independent of ε, β, u, and u_h.

Let $t \in (0, T]$ be fixed. It is easy to see that the bilinear form $B(t; \cdot, \cdot)$ is coercive on $H_0^1(\Omega)$, viz.,

$$B(t; v, v) \geq |||v|||^2, \tag{4.2}$$

for all $v \in H_0^1(\Omega)$, and is continuous in the following sense:

$$B(t; w, v) \lesssim (|||w||| + |w|_A)|||v|||, \tag{4.3}$$

for all $w \in H_0^1(\Omega) + V_h$ and $v \in H_0^1(\Omega)$. Moreover, the discrete bilinear form is coercive for v_h in V_h with respect to the $||| \cdot |||$ norm, viz.,

$$B(t; v_h, v_h) + K_h(v_h, v_h) \gtrsim |||v_h|||^2 \tag{4.4}$$

Next, we introduce the following notation which will be used to define the a posteriori estimators:

$$\alpha_K := \min(h_K \varepsilon^{-\frac{1}{2}}, \beta^{-\frac{1}{2}}), \quad \alpha_E := \min(h_E \varepsilon^{-\frac{1}{2}}, \beta^{-\frac{1}{2}}), \quad \alpha_T = \min(\varepsilon^{-\frac{1}{2}}, \beta^{-\frac{1}{2}}). \tag{4.5}$$

An a posteriori estimator for the stationary problem inspired by [23] will be utilised in our analysis. More specifically, we have the following result whose proof is completely analogous to the one of Theorem 3.2 in [23] and is therefore omitted for brevity.

THEOREM 4.1. *For a given $t \in (0, T]$, let u^s be such that*

$$B(t; u^s, v) = (f, v) \qquad \forall v \in H_0^1(\Omega),$$

and consider $u_h^s \in V_h$ such that

$$B(t; u_h^s, v_h) + K_h(u_h^s, v_h) = (f, v_h) \qquad \forall v_h \in V_h.$$

Then the following a posteriori bound holds:

$$(|||u^s - u_h^s||| + |u^s - u_h^s|_A)^2 \lesssim \sum_{K \in \zeta} \alpha_K^2 ||f + \epsilon \Delta u_h^s - \mathbf{a} \cdot \nabla u_h^s - b u_h^s||_{L^2(K)}^2$$

$$+ \sum_{E \in \mathcal{E}^{int}(\zeta)} \epsilon^{\frac{3}{2}} \alpha_E ||[\nabla u_h^s]||_{L^2(E)}^2 + \sum_{E \in \mathcal{E}(\zeta)} \left(\frac{\gamma \epsilon}{h_E} + \beta h_E + \frac{h_E}{\epsilon} \right) ||[u_h^s]||_{L^2(E)}^2.$$

5. An a posteriori bound for the semi-discrete method. To highlight the main ideas, we begin by deriving an a posteriori bound for the semi-discrete problem.

DEFINITION 5.1. *For each $t \in (0, T]$, we define the elliptic reconstruction $w \in H_0^1(\Omega)$ to be the (unique) solution of the problem*

$$B(t; w, v) = (f - \frac{\partial u_h}{\partial t}, v) \qquad \forall v \in H_0^1(\Omega).$$

Remark The dG discretisation of Definition 5.1 is to find a function $w_h \in C^{0,1}(0, T, V_h)$ such that for each $t \in (0, T]$ we have

$$B(t; w_h, v_h) + K_h(w_h, v_h) = (f - \frac{\partial u_h}{\partial t}, v_h) \qquad \forall v_h \in V_h.$$

This, in conjunction with (3.5) and (4.4), implies that $w_h = u_h$. Thus, $|||w - u_h||| + |w - u_h|_A$ can be estimated using Theorem 4.1.

We decompose the error as follows:

$$e = u - u_h = \rho + \theta, \quad \text{with} \quad \rho := u - w \quad \text{and} \quad \theta := w - u_h. \tag{5.1}$$

The dG solution, u_h, admits a decomposition into a conforming part $u_{h,c} \in H_0^1(\Omega) \cap V_h$ and a non-conforming part $u_{h,d} \in V_h$ with $u_h = u_{h,c} + u_{h,d}$, such that:

$$|||u_{h,d}|||^2 + |u_{h,d}|_A^2 \lesssim \sum_{E \in \mathcal{E}(\zeta)} \left(\frac{\varepsilon \gamma}{h_E} + \beta h_E + \frac{h_E}{\varepsilon} \right) ||[u_h]||_{L^2(E)}^2,$$

$$||\frac{\partial u_{h,d}}{\partial t}||^2 \lesssim \sum_{E \in \mathcal{E}(\zeta)} h_E ||[\frac{\partial u_h}{\partial t}]||_{L^2(E)}^2. \tag{5.2}$$

We refer to [23] for proof of these estimates which are based on respective constructions by [16] and [29]. We further define $e_c := u - u_{h,c}$ and $\theta_c := w - u_{h,c}$.

LEMMA 5.2. *For each $t \in (0, T]$ and for all $v \in H_0^1(\Omega)$ we have*

$$(\frac{\partial e}{\partial t}, v) + B(t; \rho, v) = 0.$$

Proof. This follows directly from Definition 5.1 and (3.8). □

We define the *initial condition estimator*, $\tilde{\eta}_I$, by

$$\tilde{\eta}_I^2 = ||u_0 - u_h(0)||^2 + \sum_{E \in \mathcal{E}(\zeta)} h_E ||[u_h(0)]||_{L^2(E)}^2, \tag{5.3}$$

while the *spatial estimator*, $\tilde{\eta}_S$, is given by

$$\tilde{\eta}_S^2 = \int_0^T \tilde{\eta}_{S_1}^2 \, dt + \min \left\{ \left(\int_0^T \tilde{\eta}_{S_2} \, dt \right)^2, \alpha_T^2 \int_0^T \tilde{\eta}_{S_2}^2 \, dt \right\}, \qquad (5.4)$$

where

$$\tilde{\eta}_{S_1}^2 = \sum_{K \in \zeta} \alpha_K^2 \| f - \frac{\partial u_h}{\partial t} + \varepsilon \Delta u_h - \mathbf{a} \cdot \nabla u_h - b u_h \|_{L^2(K)}^2 + \sum_{E \in \mathcal{E}^{int}(\zeta)} \varepsilon^{\frac{3}{2}} \alpha_E \| [\nabla u_h] \|_{L^2(E)}^2$$

$$+ \sum_{E \in \mathcal{E}(\zeta)} \left(\frac{\varepsilon \gamma}{h_E} + \beta h_E + \frac{h_E}{\varepsilon} \right) \| [u_h] \|_{L^2(E)}^2,$$

$$(5.5)$$

is the residual term and

$$\tilde{\eta}_{S_2}^2 = \sum_{E \in \mathcal{E}(\zeta)} h_E \| [\frac{\partial u_h}{\partial t}] \|_{L^2(E)}^2, \qquad (5.6)$$

is the rate of change of non-conformity.

THEOREM 5.3. *The error e of the semi-discrete dG method (3.5) satisfies the bound*

$$\int_0^T \| |e| \|^2 \, dt \lesssim \tilde{\eta}_I^2 + \tilde{\eta}_S^2.$$

Proof. Choosing $v = e_c$ in Lemma 5.2 gives

$$(\frac{\partial e_c}{\partial t}, e_c) + B(t; e_c, e_c) = (\frac{\partial u_{h,d}}{\partial t}, e_c) + B(t; \theta_c, e_c). \qquad (5.7)$$

Using (4.2), (4.3), and Young's inequality we arrive to

$$\frac{d}{dt}(\|e_c\|^2) + \| |e_c| \|^2 \lesssim (\| |\theta_c| \| + |\theta_c|_A)^2 + \| \frac{\partial u_{h,d}}{\partial t} \| \|e_c\|. \qquad (5.8)$$

Let $E_c := \|e_c(T_0)\|$ with $T_0 \in [0, T]$ such that $\|e_c(T_0)\| = \max_{0 \le t \le T} \|e_c(t)\|$ then integrating on $[0, T]$ gives

$$\int_0^T \| |e_c| \|^2 \, dt \lesssim \|e_c(0)\|^2 + \int_0^T (\| |\theta_c| \| + |\theta_c|_A)^2 \, dt + E_c \int_0^T \| \frac{\partial u_{h,d}}{\partial t} \| \, dt. \qquad (5.9)$$

Integrating (5.8) on $[0, T_0]$ and using Young's inequality once again gives

$$(E_c)^2 \lesssim \|e_c(0)\|^2 + \int_0^T (\| |\theta_c| \| + |\theta_c|_A)^2 \, dt + \left(\int_0^T \| \frac{\partial u_{h,d}}{\partial t} \| \, dt \right)^2. \qquad (5.10)$$

Using (5.10) to bound E_c on the right-hand side of (5.9) yields

$$\int_0^T \| |e_c| \|^2 \, dt \lesssim \|e_c(0)\|^2 + \int_0^T (\| |\theta_c| \| + |\theta_c|_A)^2 \, dt + \left(\int_0^T \| \frac{\partial u_{h,d}}{\partial t} \| \, dt \right)^2. \qquad (5.11)$$

Going back to (5.8), using the Poincaré-Friedrichs inequality on $\|e_c\|$ and working as above gives

$$\int_0^T \| |e_c| \|^2 \, dt \lesssim \|e_c(0)\|^2 + \int_0^T (\| |\theta_c| \| + |\theta_c|_A)^2 \, dt + \alpha_T^2 \int_0^T \| \frac{\partial u_{h,d}}{\partial t} \|^2 \, dt. \qquad (5.12)$$

Using the triangle inequality, we have

$$|||e||| \le |||e_c||| + |||u_{h,d}||| \quad \text{and} \quad |||\theta_c||| + |\theta_c|_A \le |||\theta||| + |\theta|_A + |||u_{h,d}||| + |u_{h,d}|_A. \tag{5.13}$$

Combining (5.11), (5.12), and (5.13) yields

$$\int_0^T |||e|||^2 \, dt \lesssim ||e_c(0)||^2 + \int_0^T (|||\theta||| + |\theta|_A)^2 \, dt + \int_0^T (|||u_{h,d}|||^2 + |u_{h,d}|_A^2) \, dt$$
$$+ \min \left\{ \left(\int_0^T ||\frac{\partial u_{h,d}}{\partial t}|| \, dt \right)^2, \alpha_T^2 \int_0^T ||\frac{\partial u_{h,d}}{\partial t}||^2 \, dt \right\}. \tag{5.14}$$

The proof then follows directly from Theorem 4.1 and (5.2).
□

6. An a posteriori bound for the fully-discrete method. We now continue by applying to the fully-discrete setting the general framework presented in the previous section.

DEFINITION 6.1. *We define the* elliptic reconstruction $w^{k+1} \in H_0^1(\Omega)$ *to be the unique solution of the elliptic problem*

$$B(t^{k+1}; w^{k+1}, v) = (f^{k+1} - \frac{u_h^{k+1} - u_h^k}{\tau_{k+1}}, v) \qquad \forall v \in H_0^1(\Omega).$$

At each time step k, we decompose the dG solution u_h^k into a conforming part $u_{h,c}^k \in H_0^1(\Omega) \cap V_h^k$ and a non-conforming part $u_{h,d}^k \in V_h^k$ such that $u_h^k = u_{h,c}^k + u_{h,d}^k$. Given $t \in (t^k, t^{k+1}]$, we (re)define $u_h(t)$ to be the linear interpolant with respect to t of the values u_h^k and u_h^{k+1}, viz.,

$$u_h(t) := l_k(t)u_h^k + l_{k+1}(t)u_h^{k+1}, \tag{6.1}$$

where $\{l_k, l_{k+1}\}$ denotes the standard linear Lagrange interpolation basis defined on the interval $[t^k, t^{k+1}]$. We define $u_{h,c}(t)$ and $u_{h,d}(t)$ analogously. We can then decompose the error $e = u - u_h = e_c - u_{h,d}$ where $e_c = u - u_{h,c}$. It will also be useful to define $\theta^k = w^k - u_h^k = \theta_c^k - u_{h,d}^k$.

LEMMA 6.2. *Given* $t \in (t^k, t^{k+1}]$ *we have*

$$(\frac{\partial e}{\partial t}, v) + B(t; e_c, v) = B(t^{k+1}; w^{k+1}, v) - B(t, u_{h,c}, v) + (f - f^{k+1}, v) \qquad \forall v \in H_0^1(\Omega).$$

Proof. This follows from Definition 6.1 and (3.8) and by rearranging the resulting equation. □

Before proving the a posteriori bounds for the fully-discrete method, we introduce the error estimators. We begin by defining the *initial condition estimator*, η_I, by

$$\eta_I^2 = ||u_0 - u_h^0||^2 + \sum_{E \in \mathcal{E}(\zeta^0)} h_E ||[u_h^0]||_{L^2(E)}^2, \tag{6.2}$$

while the *spatial estimator*, η_S, is given by

$$\eta_S^2 = \sum_{j=0}^{n-1} \tau_{j+1} \eta_{S_1, j+1}^2 + \min \left\{ \left(\sum_{j=0}^{n-1} \int_{t^j}^{t^{j+1}} \eta_{S_2, j+1} \, dt \right)^2, \alpha_T^2 \sum_{j=0}^{n-1} \int_{t^j}^{t^{j+1}} \eta_{S_2, j+1}^2 \, dt \right\}, \tag{6.3}$$

where

$$\eta^2_{S_1,j+1} = \sum_{K \in \zeta^j \cup \zeta^{j+1}} \alpha^2_K \left|\left| f^{j+1} - \frac{u_h^{j+1} - u_h^j}{\tau_{j+1}} + \varepsilon \Delta u_h^{j+1} - \mathbf{a}^{j+1} \cdot \nabla u_h^{j+1} - b^{j+1} u_h^{j+1} \right|\right|^2_{L^2(K)}$$

$$+ \sum_{E \in \mathcal{E}^{int}(\zeta^j \cup \zeta^{j+1})} \varepsilon^{\frac{3}{2}} \alpha_E ||[\nabla u_h^{j+1}]||^2_{L^2(E)} + \sum_{E \in \mathcal{E}(\zeta^j \cup \zeta^{j+1})} \left(\frac{\gamma \varepsilon}{h_E} + \beta h_E + \frac{h_E}{\varepsilon}\right) ||[u_h^{j+1}]||^2_{L^2(E)}$$

$$+ \sum_{E \in \mathcal{E}(\zeta^j \cup \zeta^{j+1})} \left(\frac{\gamma \varepsilon}{h_E} + \beta h_E + \frac{h_E}{\varepsilon}\right) ||[u_h^{j+1} - u_h^j]||^2_{L^2(E)},$$

$$(6.4)$$

and

$$\eta^2_{S_2,j+1} = \sum_{E \in \mathcal{E}(\zeta^j \cup \zeta^{j+1})} h_E ||[\frac{u_h^{j+1} - u_h^j}{\tau_{j+1}}]||^2_{L^2(E)}$$

$$+ \sum_{E \in \mathcal{E}^{int}(\zeta^j \cup \zeta^{j+1})} h_E^{-1} ||[l_j(t)\mathbf{a}(u_h^{j+1} - u_h^j) + (\mathbf{a}^{j+1} - \mathbf{a})u_h^{j+1}]||^2_{L^2(E)}.$$

$$(6.5)$$

The final term is the *time (or temporal) estimator*, η_T, given by

$$\eta^2_T = \frac{1}{4} \sum_{j=0}^{n-1} \tau_{j+1} \eta^2_{T_1,j+1} + \min\left\{ \left(\sum_{j=0}^{n-1} \int_{t^j}^{t^{j+1}} \eta_{T_2,j+1}\, dt\right)^2, \alpha^2_T \sum_{j=0}^{n-1} \int_{t^j}^{t^{j+1}} \eta^2_{T_2,j+1}\, dt \right\}.$$

$$(6.6)$$

where

$$\eta^2_{T_1,j+1} = \sum_{K \in \zeta^j \cup \zeta^{j+1}} \epsilon ||\nabla(u_h^{j+1} - u_h^j)||^2_{L^2(K)}, \tag{6.7}$$

and

$$\eta^2_{T_2,j+1} = \sum_{K \in \zeta^j \cup \zeta^{j+1}} ||l_j(t)(\mathbf{a} \cdot \nabla(u_h^{j+1} - u_h^j) + b(u_h^{j+1} - u_h^j)) + f - f^{j+1}$$

$$+ (\mathbf{a}^{j+1} - \mathbf{a}) \cdot \nabla u_h^{j+1} + (b^{j+1} - b)u_h^{j+1}||^2_{L^2(K)}.$$

$$(6.8)$$

THEOREM 6.3. *The error e of the fully-discrete method satisfies the bound*

$$\sum_{j=0}^{n-1} \int_{t^j}^{t^{j+1}} |||e|||^2\, dt \lesssim \eta^2_I + \eta^2_S + \eta^2_T. \tag{6.9}$$

Proof. From Lemma 6.2 we have

$$(\frac{\partial e_c}{\partial t}, v) + B(t; e_c, v) = (\frac{\partial u_{h,d}}{\partial t}, v) + B(t^{k+1}; w^{k+1}, v) - B(t, u_{h,c}, v) + (f - f^{k+1}, v),$$

$$(6.10)$$

which upon straightforward manipulation and setting $v = e_c$ gives

$$\frac{d}{dt}(||e_c||^2) + B(t; e_c, e_c) = (\frac{\partial u_{h,d}}{\partial t}, e_c) + B(t^{k+1}; \theta_c^{k+1}, e_c) + l_k(t)B(t; u_{h,c}^{k+1} - u_{h,c}^k, e_c)$$

$$+ B(t^{k+1}; u_{h,c}^{k+1}, e_c) - B(t; u_{h,c}^{k+1}, e_c) + (f - f^{k+1}, e_c).$$

$$(6.11)$$

9

Denoting by $I_h^k e_c$ the L^2-projection of e_c onto $H_0^1(\Omega) \cap V_h(\zeta^k \cup \zeta^{k+1})$, using (4.2), (4.3), the Cauchy-Schwarz inequality, Young's inequality and the triangle inequality yields

$$\frac{d}{dt}(||e_c||^2) + |||e_c|||^2 \lesssim ||\frac{\partial u_{h,d}}{\partial t}||||e_c|| + (|||\theta^{k+1}||| + |\theta^{k+1}|_A)^2 + |||u_{h,d}|||^2 + |u_{h,d}|_A^2$$
$$+ T_1 + T_2 + T_3 + T_4 + T_5, \tag{6.12}$$

where

$$T_1 := \sum_{K \in \zeta^k \cup \zeta^{k+1}} \int_K (f - f^{k+1} + (\mathbf{a}^{k+1} - \mathbf{a}) \cdot \nabla u_h^{k+1} + (b^{k+1} - b)u_h^{k+1}$$
$$+ l_k(t)(\mathbf{a} \cdot \nabla(u_h^{k+1} - u_h^k) + b(u_h^{k+1} - u_h^k)))e_c \, dx, \tag{6.13}$$

$$T_2 := l_k(t) \sum_{K \in \zeta^k \cup \zeta^{k+1}} \int_K \epsilon \nabla(u_h^{k+1} - u_h^k) \cdot \nabla e_c \, dx, \tag{6.14}$$

$$T_3 := \frac{1}{2} \sum_{K \in \zeta^k \cup \zeta^{k+1}} \int_{\partial K \backslash \partial \Omega} [(\mathbf{a}^{k+1} - \mathbf{a})u_h^{k+1}](e_c - I_h^k e_c) \, ds, \tag{6.15}$$

$$T_4 := \frac{1}{2} l_k(t) \sum_{K \in \zeta^k \cup \zeta^{k+1}} \int_{\partial K \backslash \partial \Omega} [\mathbf{a}(u_h^{k+1} - u_h^k)](e_c - I_h^k e_c) \, ds, \tag{6.16}$$

and

$$T_5 := \frac{1}{2} \sum_{K \in \zeta^k \cup \zeta^{k+1}} \int_{\partial K \backslash \partial \Omega} [l_k(t)\mathbf{a}(u_h^{k+1} - u_h^k) + (\mathbf{a}^{k+1} - \mathbf{a})u_h^{k+1}] I_h^k e_c \, ds. \tag{6.17}$$

Using the discrete and integral versions of the Cauchy-Schwarz inequality, we immediately have

$$T_1 \le \eta_{T_2,k+1} ||e_c||, \qquad T_2 \le \eta_{T_1,k+1} |||e_c|||. \tag{6.18}$$

Further, using the approximation properties of I_h^k we have

$$T_3 + T_4 \lesssim \eta_{S_1,k+1} |||e_c|||. \tag{6.19}$$

Next, using the standard inverse estimate $||I_h^k e_c||_{L^2(E)}^2 \lesssim h_E^{-1} ||I_h^k e_c||_{L^2(K)}^2$, for $E \subset \partial K$, and the stability of the L^2-projection $||I_h^k e_c|| \le ||e_c||$, we deduce

$$T_5 \lesssim \eta_{S_2,k+1} ||e_c||. \tag{6.20}$$

The proof then follows from Theorem 4.1 and (5.2) and by employing a bounding strategy identical to that used in Theorem 5.3. □

RemarkAn alternative a posteriori bound can be show, using directly (4.3) on (6.11). The resulting bound will contain computable terms in the $|\cdot|_A$-norm, which can, in turn, be estimated by a more easily localisable term of the form $\epsilon^{-1/2}||\cdot||_{L^2(\Omega)}$.

Numerical experiments (omitted here for brevity) show that this alternative estimator produces very similar results to the ones based on Theorem 6.3.

RemarkAlthough efficiency bounds are not presented in this work, we refer to [28] for the proof of efficiency bounds for the a posteriori estimator presented therein for the SUPG spatial setting which should be extendable to the estimator presented in Theorem 6.3 given the similarity of some of the terms presented.

RemarkThe use of elliptic reconstruction is not essential for the proof of Theorem 5.3 and Theorem 6.3. It is possible to derive the residual based a posteriori bounds directly albeit at the cost of a lengthier calculation. The advantage of using the elliptic reconstruction in the proof lies in the fact that it can be easily modified to include non-residual based spatial a posteriori estimators. This, in turn, may offer improvements in robustness with respect to the Péclet number (cf. [21]).

7. An adaptive algorithm. The a posteriori bounds presented above will be used to drive a space-time adaptive algorithm. A number of adaptive algorithms for parabolic problems have been proposed in the literature; see e.g. [8, 22] and the references therein.

Here, we propose a variant of the adaptive algorithms from [8, 22] which appears to perform well for our discretisation. The pseudocode is given in Algorithm 1. The algorithm is based on using different parts of the a posteriori estimator from Theorem (6.3) to drive space-time adaptivity.

Both mesh refinement and coarsening are driven by the term $\eta_{S_1,j+1}$. The size of the elemental contributions to $\eta_{S_1,j+1}$ determines the percentages of elements to be refined and/or coarsened depending on the two spatial tolerances `stola` and `stolb`. The nature of the time estimator, η_T, makes it difficult to use as a temporal refinement indicator so, to this end we define $\hat{\eta}_{T,j+1}$ given by

$$\hat{\eta}_{T,j+1}^2 = \frac{1}{4}\tau_{j+1}\eta_{T_1,j+1}^2 + \min\{\alpha_T, T\}\int_{t^j}^{t^{j+1}} \eta_{T_2,j+1}^2 \, dt. \tag{7.1}$$

The sum of all terms $\hat{\eta}_{T,j+1}$ bound η_T and hence can be used to drive temporal refinement subject to a time tolerance `ttol` on each time interval.

RemarkThe term $\eta_{S_2,j+1}$ is smaller than $\eta_{S_1,j+1}$ both from a theoretical and a practical standpoint. It is, therefore, not used to drive the spatial mesh modification.

A difference of Algorithm 1 to (standard) adaptivity algorithms for conforming finite elements for parabolic problems is that here a limit is given on how much the mesh can change. This is introduced in order to to ensure termination of the algorithm. The input value m is a bound on how many times the mesh can change per unit time. The value of m can be very large in practice.

8. Numerical experiments. We shall investigate numerically the presented a posteriori bounds and the performance of the adaptive algorithm through an implementation based on the `deal.II` finite element library ([4]). All the numerical experiments have been performed using the high performance computing facility ALICE at the University of Leicester.

We denote by λ_k the total number of degrees of freedom on the union mesh $\zeta^k \cup \zeta^{k+1}$. Hence, the weighted total number of degrees of freedom of the problem is given by

$$Total\ DoFs := \sum_{j=0}^{n-1} \tau_{j+1}\lambda_j. \tag{8.1}$$

11

Algorithm 1 Space-time adaptivity

1: **Input:** $\epsilon, \mathbf{a}, b, f, u_0, T, \Omega, m, n, \zeta^0, \gamma, \mathtt{initol}, \mathtt{ttol}, \mathtt{stola}, \mathtt{stolb}, \mathtt{ref}\%, \mathtt{coar}\%$.
2: Calculate u_h^0.
3: Set $j = 0$, $\tau_1, ..., \tau_n = T/n$, $time = \tau_1$, $threshold = T/m$, $counter = 0$.
4: **while** $\eta_I > initol$ **do**
5:     Refine ζ^0 by 10% and coarsen ζ^0 by 5%.
6:     Calculate u_h^0.
7: **end while**
8: **while** $time \leq T$ **do**
9:     Calculate u_h^{j+1} from u_h^j.
10:     **while** $\hat{\eta}_{T,j+1} > \mathtt{ttol}$ **do**
11:         $n \leftarrow n + 1$
12:         $time \leftarrow time - \tau_{j+1}$
13:         $\tau_{j+3} = \tau_{j+2}, ..., \tau_n = \tau_{n-1}$
14:         $\tau_{j+2} = \tau_{j+1}/2$
15:         $\tau_{j+1} \leftarrow \tau_{j+1}/2$
16:         $time + \tau_{j+1}$
17:         Calculate u_h^{j+1} from u_h^j.
18:     **end while**
19:     $counter \leftarrow counter + \tau_{j+1}$
20:     **if** $\eta_{S_1,j+1} > \mathtt{stola}$ and $counter > threshold$ **then**
21:         Create ζ^{j+1} from ζ^j by refining by $\mathtt{ref}\%$.
22:     **end if**
23:     **if** $\mathtt{stolb} < \eta_{S_1,j+1} < \mathtt{stola}$ and $counter > threshold$ **then**
24:         Create ζ^{j+1} from ζ^j by refining by $\mathtt{ref}\%$ and coarsening by $\mathtt{coar}\%$.
25:     **end if**
26:     **if** $\eta_{S_1,j+1} < \mathtt{stolb}$ and $counter > threshold$ **then**
27:         Create ζ^{j+1} from ζ^j by coarsening by $\mathtt{coar}\%$.
28:     **end if**
29:     Calculate u_h^{j+1} from u_h^j.
30:     $j \leftarrow j + 1$
31:     $time \leftarrow time + \tau_{j+1}$
32:     **if** $counter < threshold$ **then**
33:         $counter = 0$
34:     **end if**
35: **end while**

In all examples presented below, unless otherwise stated, $\gamma = 10$, $\mathtt{stolb} = \mathtt{stola}/5$, $\mathtt{ref}\% = 6.25\%$ and ζ^0 is an 8×8 uniform quadrilateral mesh. We further take $\mathtt{coar}\% = 10\%$ for Example 1, Example 2 and Example 4 while we set $\mathtt{coar}\% = 30\%$ for Example 3.

    8.1. Example 1. Let $\Omega = (0,1)^2$, $\mathbf{a} = (1,1)^T$, $b = 0$, $u_0 = 0$, $T = 10$ and select the function f so that the exact solution to problem (2.5) is given by

$$u(x,y,t) = (1 - e^{-t})\left(\frac{e^{(x-1)/\varepsilon} - 1}{e^{-1/\varepsilon} - 1} + x - 1\right)\left(\frac{e^{(y-1)/\varepsilon} - 1}{e^{-1/\varepsilon} - 1} + y - 1\right). \qquad (8.2)$$

The solution exhibits boundary layers at the outflow boundary of the domain of width $\mathcal{O}(\varepsilon)$ as well as a temporal boundary layer.

Time steps	Total DoFs	Est.	Est.Ratio	Error	Err.Ratio
10	160	3.83e-1		6.45e-2	
20	640	1.94e-1	0.507	3.07e-2	0.475
40	2560	9.79e-2	0.502	1.47e-2	0.479
80	10240	4.90e-2	0.501	7.17e-3	0.487
160	40960	2.45e-2	0.500	3.53e-3	0.492
320	163840	1.22e-2	0.500	1.75e-3	0.496
640	655360	6.14e-3	0.500	8.72e-4	0.497
1280	2621440	3.07e-3	0.500	4.35e-4	0.498

TABLE 8.1

Example 1. Decay of the a posteriori error estimator for $p = 1$, $\varepsilon = 1$.

Time steps	Total DoFs	Est.	Est.Ratio	Error	Err.Ratio
10	160	1.66e+1		8.56e-1	
20	640	1.29e+1	0.773	1.06e+0	1.240
40	2560	1.08e+1	0.838	1.18e+0	1.114
80	10240	8.68e+0	0.802	1.27e+0	1.078
160	40960	5.91e+0	0.680	1.14e+0	0.899
320	163840	3.47e+0	0.587	7.54e-1	0.657
640	655360	1.88e+0	0.542	4.06e-1	0.538
1280	2621440	0.98e+0	0.520	2.01e-1	0.496

TABLE 8.2

Example 1. Decay of the a posteriori error estimator for $p = 1$, $\varepsilon = 10^{-2}$.

We begin by assessing the decay of the estimators on uniform spatial and temporal meshes. To this end, the number of time-steps is doubled while the number of spatial degrees of freedom is quadrupled by respective uniform refinements. The results for the spatial polynomial degree $p = 1$ are given in Tables 8.1 and 8.2. As expected, the error halves (Err. Ratio) for each space-time uniform refinement and the same is observed for the a posteriori estimator (Est. Ratio), implying optimality of the a posteriori error estimator for this example. This is achieved immediately for $\varepsilon = 1$ and upon sufficient resolution of the boundary layer for $\varepsilon = 10^{-2}$. The a posteriori estimator appears to be of optimal rate under uniform refinement in all the numerical examples presented in this work; these results are omitted for brevity.

We now assess the performance of the space-time adaptive algorithm presented in Section 7; we do this for $p > 1$ as the case $p = 1$ needs special care and hence will be discussed separately.

We begin by fixing a temporal tolerance that produces enough time steps so that the temporal contribution to the error is very small in comparison to the spatial contribution. The spatial tolerance is then gradually reduced to compare the rates of convergence of the spatial estimator with that of the actual total error. Results for $\epsilon = 1$ and $\epsilon = 10^{-2}$ and for polynomial degrees $p = 2$ and $p = 3$ are shown in Figure 8.1; here as in all the following convergence plots an unmarked line indicates the optimal order of convergence. Optimal rates of convergence are observed with respect to the total degrees of freedom for both the estimator and the error. Further, as shown in the bottom plots of Figure 8.1, the effectivity indices are bounded and remain between 6 and 11 for the two values of ϵ; these are directly comparable to those observed in [23] for the stationary problem.
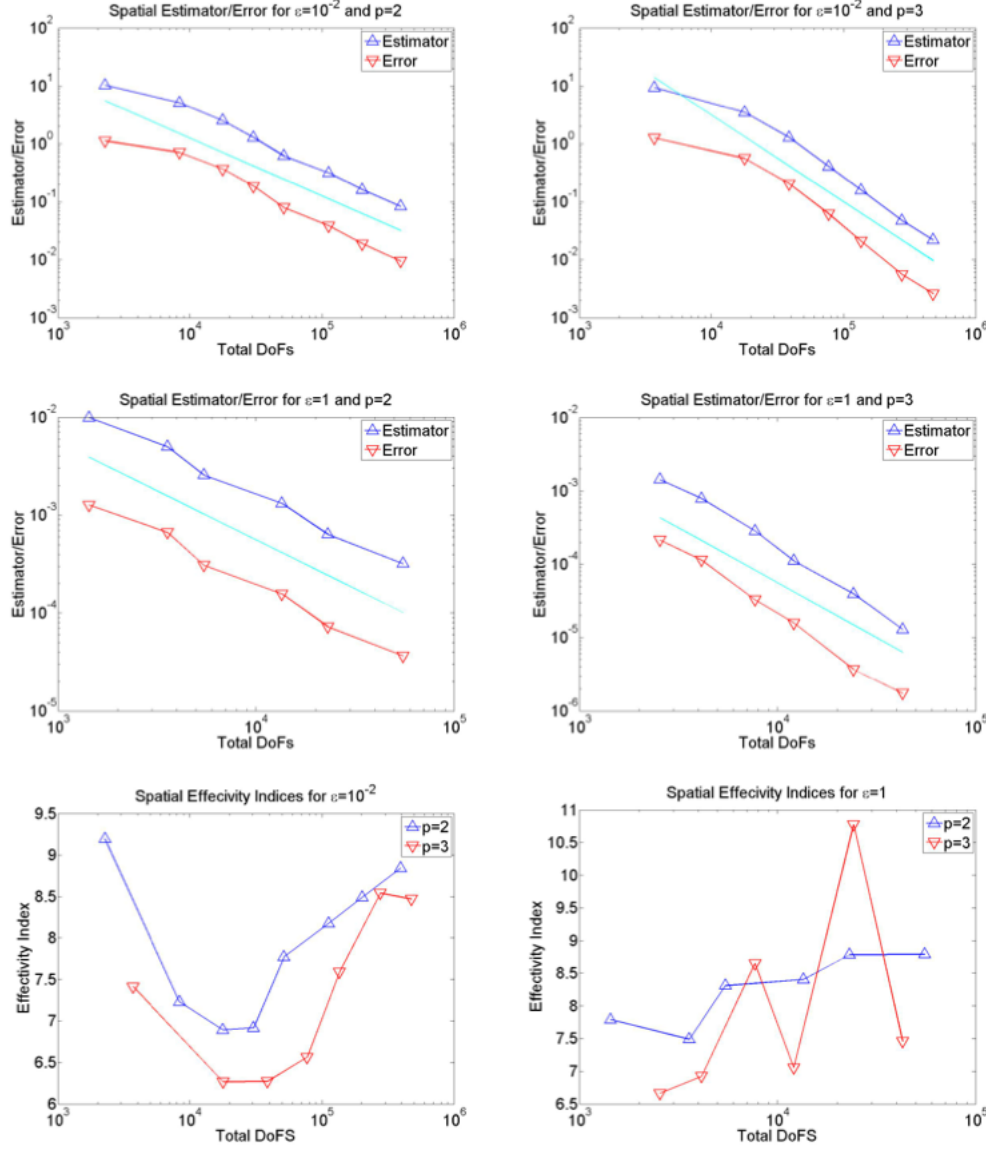
13

FIG. 8.1. *Example 1: Spatial rates of convergence and effectivity indices for $\epsilon = 1, \epsilon = 10^{-2}$ and $p = 2, 3$*

In order to study the temporal rates of convergence of the adaptive algorithm any boundary layers must be fully resolved so that the spatial error is dominated by the temporal one. To this end, we set $\gamma = 100$ and use $p = 10$ on a sufficiently fine fixed spatial mesh. We monitor the algorithm's behaviour upon gradually reducing the temporal tolerance. The results are given in Figure 8.2. Optimal order is observed with respect to both the time estimator and the error. The effectivity indices appear to remain bounded and converge to a smaller value for $\epsilon = 10^{-2}$ than for $\epsilon = 1$ which is evidence of robustness with respect to the Péclet number.

The presence of a temporal boundary layer in the solution motivates a comparison
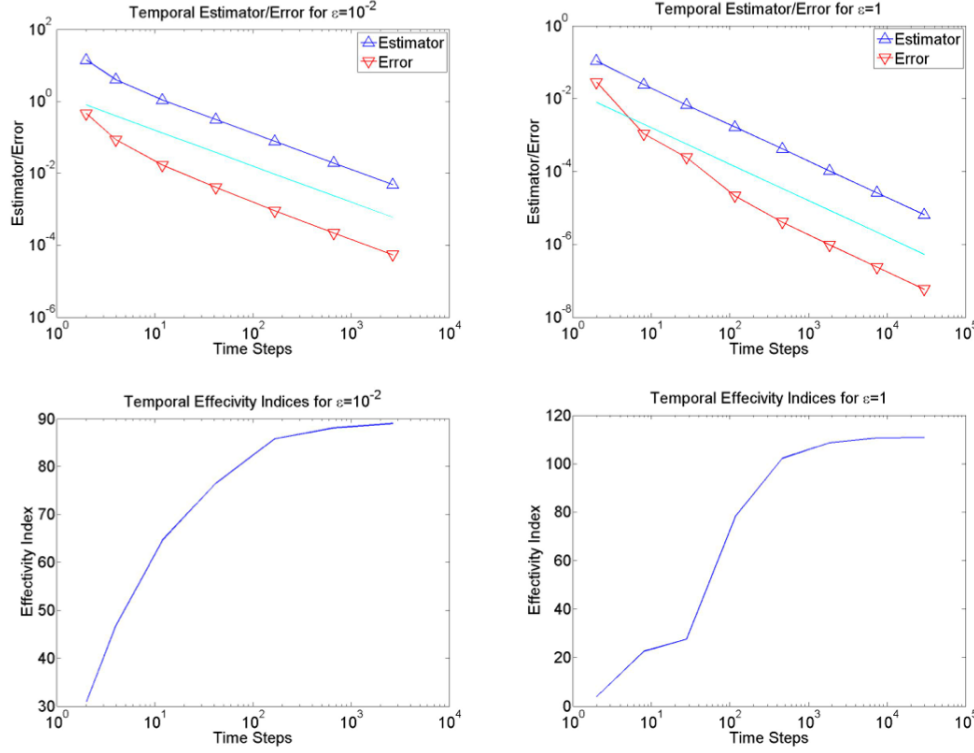
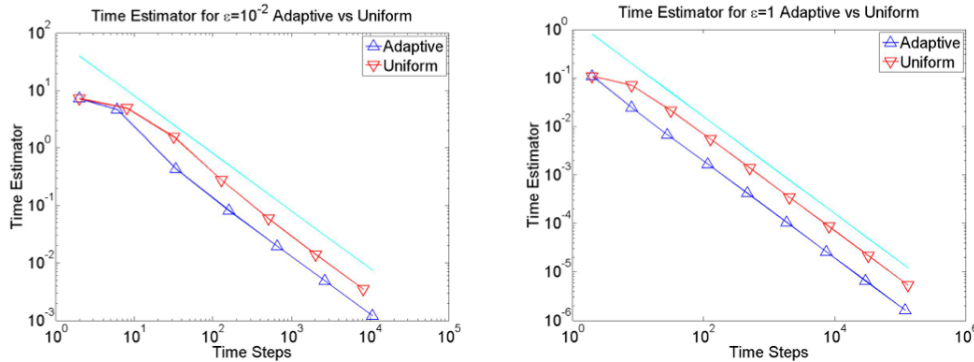FIG. 8.2. *Example 1: Temporal rates of convergence and effectivity indices for* $\epsilon = 1, 10^{-2}$.

FIG. 8.3. *Example 1: Adaptive vs uniform comparison for* $\epsilon = 1, 10^{-2}$.

between adaptive and uniform time-stepping. We do this by using $p = 3$ and a sufficiently resolved spatial mesh and by reducing the temporal tolerance in an adaptive fashion and uniformly, respectively. The results are given in the plots in Figure 8.3: the savings achieved by using adaptive time-stepping are evident.

Let us now discuss the behaviour of the space-time a posteriori estimator in the case $p = 1$. Recall that the estimator has already been proven to be of optimal rate in both space and time under uniform refinement, *cf.* Tables 8.1 and 8.2. In the adaptive setting, however, where the mesh can change between time steps, we observe that the
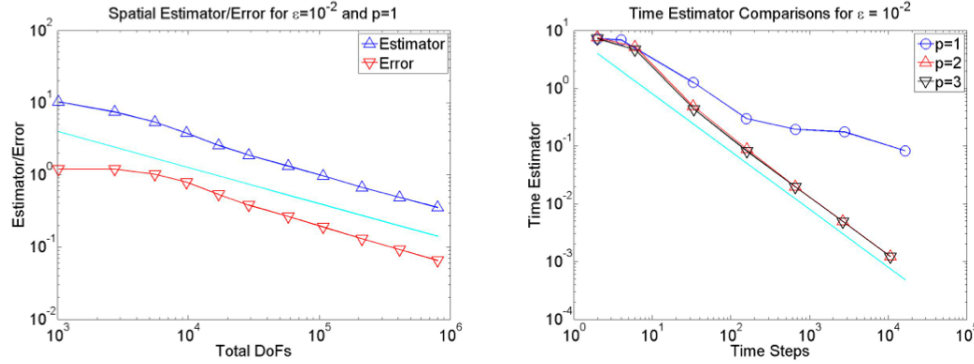
FIG. 8.4. *Example 1: Spatial and temporal rates of convergence for $\epsilon = 10^{-2}$ and $p = 1$.*

rate of convergence of the time estimator for $p = 1$ may be slower than the expected optimal rate. This is indeed the case for this test problem, as shown in the right plot of Figure 8.4 by comparison with the results obtained with $p = 2$ and $p = 3$. On the other hand, the spatial estimator retains optimal rate, as shown in the left plot of Figure 8.4.

Notice that the method itself could be suboptimal under mesh change for $p = 1$; this is something we have not been able to test and indeed the nature of this degeneracy remains unclear to the authors. It is also unclear if this is, or not, a special feature of the backward-Euler-dG method considered in this work. In the literature, numerical experiments including *rates* of convergence of space-time adaptive algorithms for parabolic problems are very rare, especially for non-conforming methods. Therefore, it is not possible at this point to assess the generality of this phenomenon. We mention in passing the recent work [5] where degeneracy due to mesh modification (including mesh refinement) has been observed for Crank-Nicolson FEM discretisations. As shown above, the degeneracy of the temporal rate of convergence when spatial mesh-change is present disappears for $p > 1$. Hence, for the remainder of this work, all the numerical experiments with space-time adaptivity are performed with $p > 1$.

8.2. Example 2. We set $\Omega = (-1, 1)^2$, $\mathbf{a} = (1, 1)^T$, $b = 1$, $f = sin(5t)xy$, $u_0 = 0$ and $T = 2\pi$. The solution exhibits layers of width $\mathcal{O}(\varepsilon)$ in the proximity of the outflow boundary and is oscillatory in time. The sharpness of the boundary layers depend on time, thus making this a good test of the ability of the algorithm to add and remove degrees of freedom.

As in Example 1, we begin by fixing a temporal tolerance while decreasing the spatial tolerance to observe the rates of convergence for the space estimator. We then set $p = 3$ with a spatial tolerance small enough to resolve any boundary layers, while reducing the temporal tolerance to observe the rates of the time estimator. The results are displayed in Figure 8.5. Optimal rates of convergence are observed for both the space and the time estimators.

To assess the mesh change driven by the adaptive algorithm we also plot the individual degrees of freedom on each mesh against time for a given spatial tolerance and temporal tolerance. The results are given in Figure 8.6. We observe that the adaptive algorithm is adding and removing degrees of freedom at a rate that is in accordance with the oscillating nature of the solution driven by the sinusoidal forcing
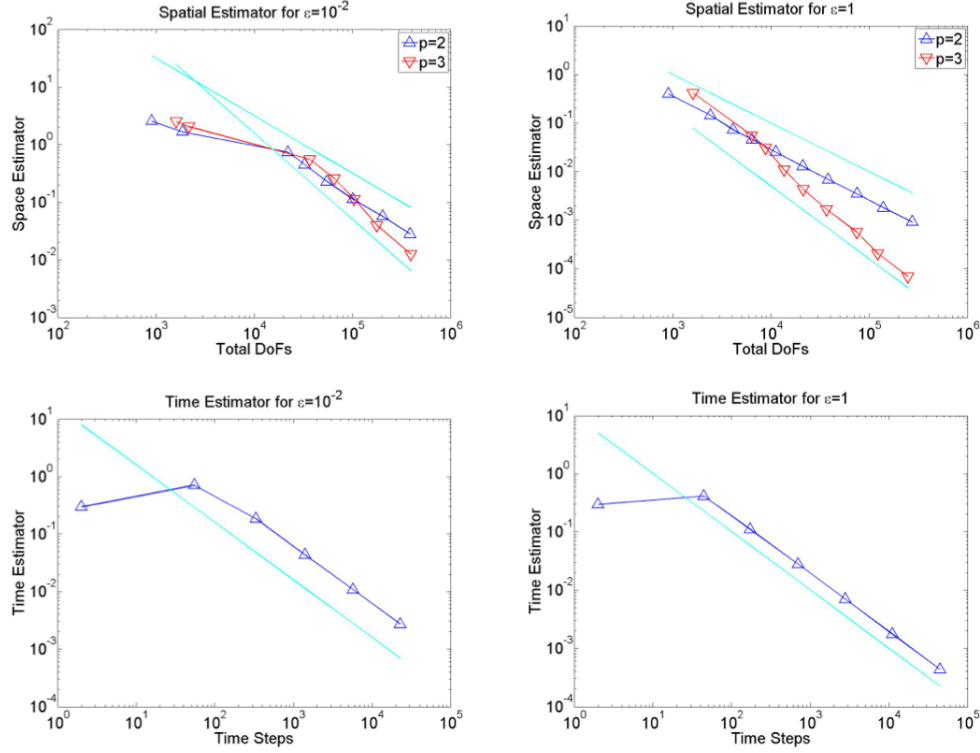
16

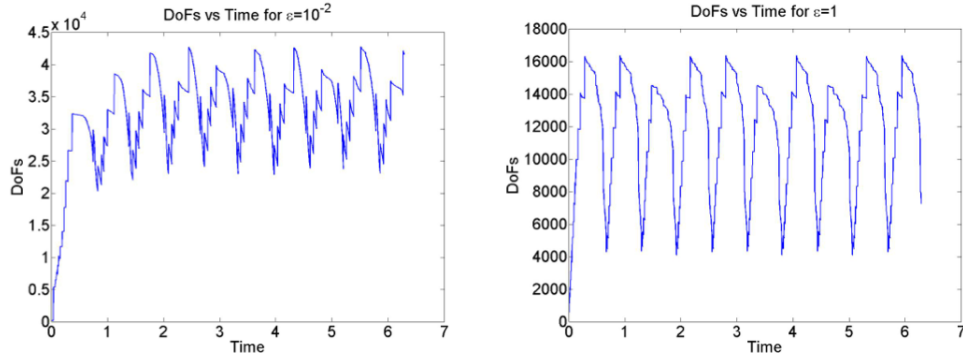FIG. 8.5. *Example 2: Spatial and temporal rates of convergence for* $\varepsilon = 1, 10^{-2}$.

FIG. 8.6. *Example 2: Degrees of freedom on each time step plotted against time for* $\varepsilon = 1, 10^{-2}$

function f.

8.3. Example 3. Let $\Omega = (-1, 1)^2$, $T = 100$, $\mathbf{a} = (y, -x)^T$, $b = 0$, $f = 0$, and $u_0 = e^{-64(x-0.5)^2} e^{-64(y-0.5)^2}$. The PDE convects the initial two dimensional Gaussian profile along the circular wind while diffusing it at a rate depending upon ε.

Setting $p = 3$ and fixing a spatial tolerance and an initial tolerance, we reduce the temporal tolerance to observe the temporal rates of the problem. Additionally, for $\epsilon = 1$, we compare adaptive time-stepping and uniform time-stepping. The results are given in Figure 8.7. Some meshes at various time steps produced by the algorithm
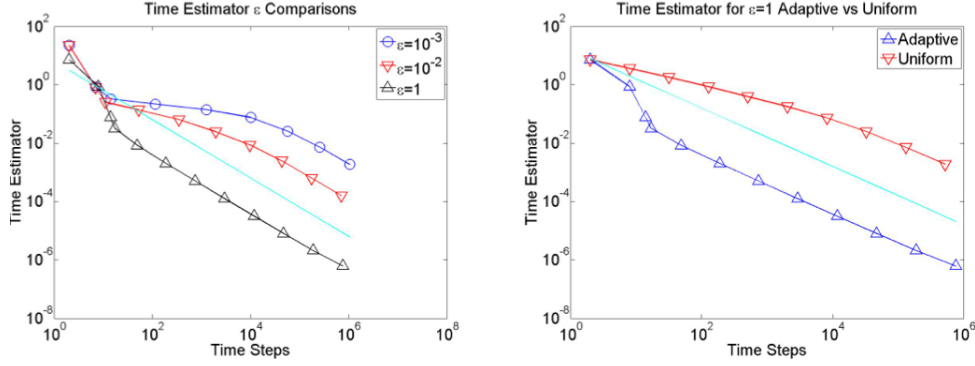
17

Fig. 8.7. *Example 3: Convergence of the time estimator for $\varepsilon = 1, 10^{-2}, 10^{-3}$ and comparison of adaptive vs. uniform time steps.*
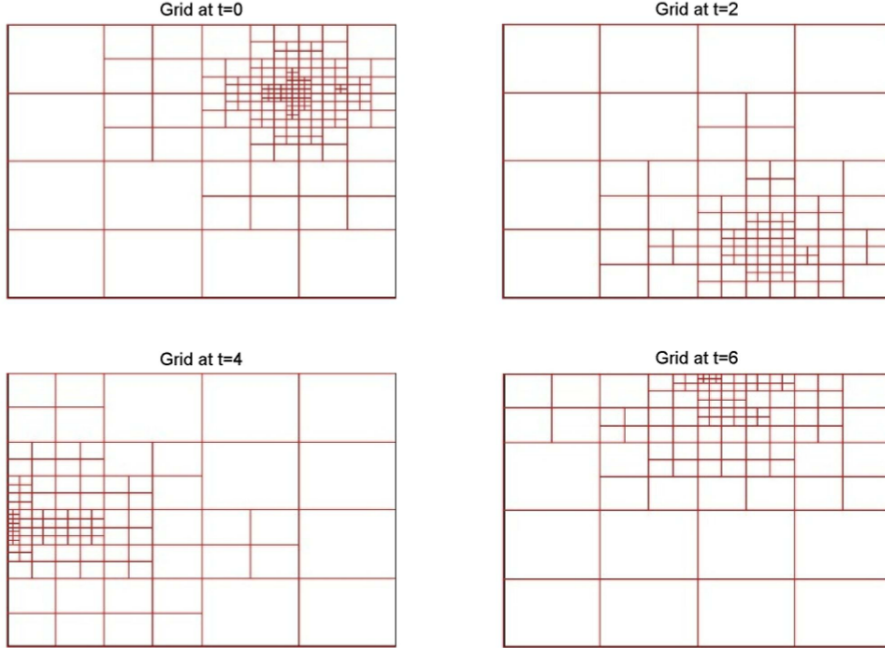

Fig. 8.8. *Example 3: grid snapshots.*

for $\epsilon = 10^{-3}$ are shown in Figure 8.8.

The smaller the value of ϵ, the longer it takes the time estimator to reach optimal order which is to be expected theoretically. Furthermore, large savings are observed using adaptive time-stepping over uniform time-stepping when the diffusion is large; time steps would be otherwise wasted when the solution is effectively zero.

8.4. Example 4. Let $\Omega = (0,1)^2$, $\mathbf{a} = (sin(t), cos(t))^T$, $b = 0$, $f = 1$, $u_0 = 0$ and $T = 2\pi$. The nature of the solution is rather uniform in time but has a boundary layer of width $\mathcal{O}(\varepsilon)$ whose location does depend on time. Therefore, this example is well suited to test the ability of the algorithm to adapt the grid to the feature of the solution as time evolves.
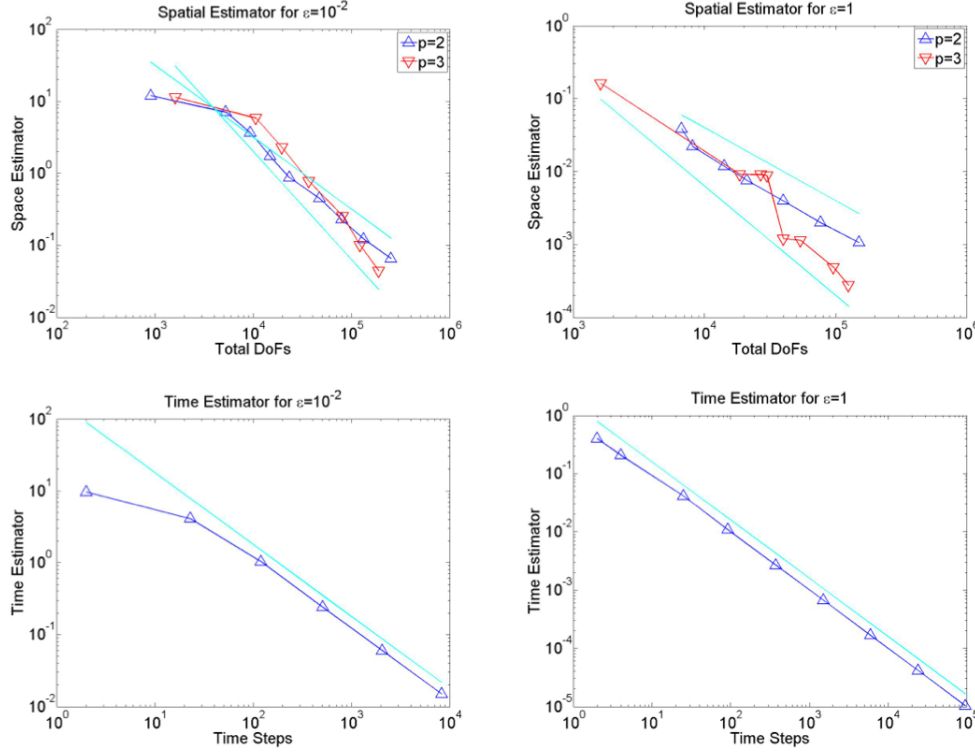
FIG. 8.9. *Example 4: Spatial and temporal rates of convergence for $\varepsilon = 1, 10^{-2}$.*

As in previous examples, we fix a temporal tolerance and then reduce the spatial tolerance to observe the rates of the space estimator. Again, we also fix $p = 3$ and a spatial tolerance small enough to ensure that all boundary layers are sufficiently resolved and then observe the rates of the time estimator. These results are given in Figure 8.9. Finally, grids at various times for $\epsilon = 10^{-2}$ and $p = 2$ are given in Figure 8.10. Optimal spatial and temporal rates of convergence are observed. The grids produced for $\epsilon = 10^{-2}$ clearly show that the adaptive algorithm is picking up the boundary layers as they move around the domain and that unneeded degrees of freedom are not retained.

9. Conclusions. An a posteriori error estimator for the discontinuous Galerkin spatial discretisation of a non-stationary linear convection-diffusion equation is presented. The derivation of the estimator is based on reconstruction techniques to make use of robust a posteriori estimators for elliptic problems already developed in the literature. Our numerical examples clearly indicate that the error estimator is robust and the respective space-time adaptive algorithm works well for the studied problems. The spatial effecitivity indices are in an identical range to those observed in ([23]) and the temporal effectivity indices are approximately of the same order as those currently seen in the literature for similar problems ([13]) and appear independent of ϵ. Further study to understand the temporal rate of convergence for $p = 1$ under mesh-modification as well as extensions to high order time-stepping methods will be considered in the future.
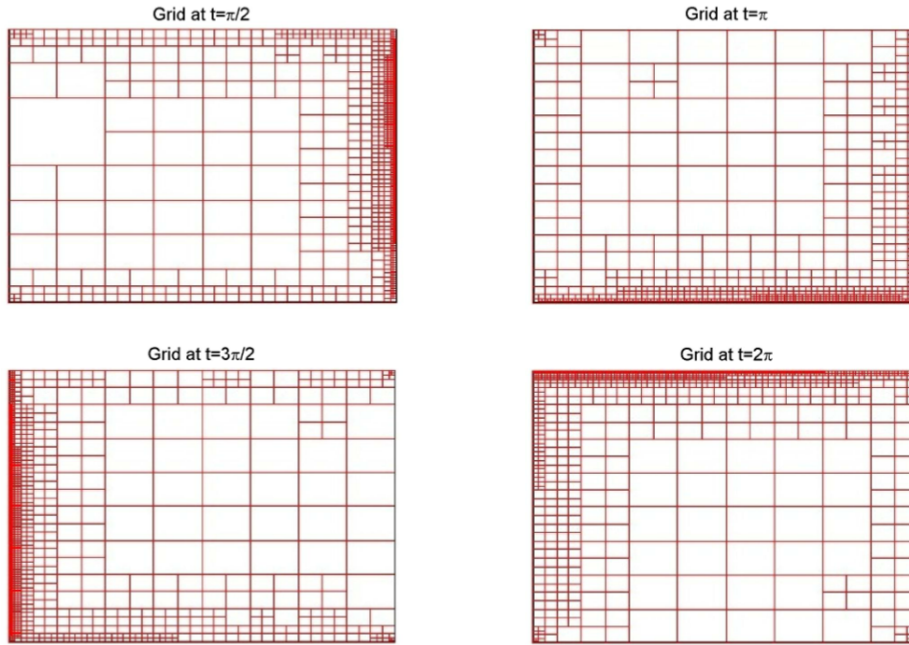
REFERENCES

Fig. 8.10. *Example 4: grid snapshots.*

[1] Mark Ainsworth and J. Tinsley Oden. *A posteriori error estimation in finite element analysis.* Pure and Applied Mathematics (New York). Wiley-Interscience [John Wiley & Sons], New York, 2000.

[2] Rodolfo Araya, Edwin Behrens, and Rodolfo Rodríguez. An adaptive stabilized finite element scheme for the advection-reaction-diffusion equation. *Appl. Numer. Math.*, 54(3-4):491–503, 2005.

[3] Rodolfo Araya, Abner H. Poza, and Ernst P. Stephan. A hierarchical a posteriori error estimate for an advection-diffusion-reaction problem. *Math. Models Methods Appl. Sci.*, 15(7):1119–1139, 2005.

[4] W. Bangerth, R. Hartmann, and G. Kanschat. deal.II—a general-purpose object-oriented finite element library. *ACM Trans. Math. Software*, 33(4):Art. 24, 27, 2007.

[5] Eberhard; Bänsch, Foteini; Karakatsani, and Charalambos Makridakis. A posteriori error control for fully discrete Crank-Nicolson schemes. *SIAM J. Numer. Anal.*, page to appear, 2012.

[6] Roland Becker, Peter Hansbo, and Mats G. Larson. Energy norm a posteriori error estimation for discontinuous Galerkin methods. *Comput. Methods Appl. Mech. Engrg.*, 192(5-6):723–733, 2003.

[7] Stefano Berrone and Claudio Canuto. Multilevel a posteriori error analysis for reaction-convection-diffusion problems. *Appl. Numer. Math.*, 50(3-4):371–394, 2004.

[8] Zhiming Chen and Jia Feng. An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems. *Math. Comp.*, 73(247):1167–1193 (electronic), 2004.

[9] Todd Dupont. Mesh modification for evolution equations. *Math. Comp.*, 39(159):85–107, 1982.

[10] Alexandre Ern and Jennifer Proft. A posteriori discontinuous Galerkin error estimates for transient convection-diffusion equations. *Appl. Math. Lett.*, 18(7):833–841, 2005.

[11] Alexandre Ern and Martin Vohralík. A posteriori error estimation based on potential and flux reconstruction for the heat equation. *SIAM J. Numer. Anal.*, 48(1):198–223, 2010.

[12] Emmanuil H. Georgoulis, Edward Hall, and Paul Houston. Discontinuous Galerkin methods for advection-diffusion-reaction problems on anisotropically refined meshes. *SIAM J. Sci. Comput.*, 30(1):246–271, 2008.

[13] Emmanuil H. Georgoulis, Omar Lakkis, and Juha M. Virtanen. A posteriori error control for discontinuous Galerkin methods for parabolic problems. *SIAM J. Numer. Anal.*, 49(2):427–458, 2011.

[14] Paul Houston, Dominik Schötzau, and Thomas P. Wihler. Energy norm a posteriori error estimation of hp-adaptive discontinuous Galerkin methods for elliptic problems. *Math. Models Methods Appl. Sci.*, 17(1):33–62, 2007.

[15] Paul Houston and Endre Süli. Adaptive Lagrange-Galerkin methods for unsteady convection-diffusion problems. *Math. Comp.*, 70(233):77–106, 2001.

[16] Ohannes A. Karakashian and Frederic Pascal. A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems. *SIAM J. Numer. Anal.*, 41(6):2374–2399 (electronic), 2003.

[17] Gerd Kunert. A posteriori error estimation for convection dominated problems on anisotropic meshes. *Math. Methods Appl. Sci.*, 26(7):589–617, 2003.

[18] Alexei Lozinski, Marco Picasso, and Virabouth Prachittham. An anisotropic error estimator for the Crank-Nicolson method: application to a parabolic problem. *SIAM J. Sci. Comput.*, 31(4):2757–2783, 2009.

[19] Charalambos Makridakis and Ricardo H. Nochetto. Elliptic reconstruction and a posteriori error estimates for parabolic problems. *SIAM J. Numer. Anal.*, 41(4):1585–1594, 2003.

[20] Hans-Görg Roos, Martin Stynes, and Lutz Tobiska. *Robust numerical methods for singularly perturbed differential equations*, volume 24 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2008. Convection-diffusion-reaction and flow problems.

[21] Giancarlo Sangalli. Robust a-posteriori estimator for advection-diffusion-reaction problems. *Math. Comp.*, 77(261):41–70 (electronic), 2008.

[22] Alfred Schmidt and Kunibert G. Siebert. *Design of adaptive finite element software*, volume 42 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, Berlin, 2005. The finite element toolbox ALBERTA, With 1 CD-ROM (Unix/Linux).

[23] Dominik Schötzau and Liang Zhu. A robust a-posteriori error estimator for discontinuous Galerkin methods for convection-diffusion equations. *Appl. Numer. Math.*, 59(9):2236–2255, 2009.

[24] Dominik Schötzau and Liang Zhu. A robust *a posteriori* error estimate for hp-adaptive DG methods for convection-diffusion equations. *IMA J. Numer. Anal.*, 31(3):971–1005, 2011.

[25] Ivana Šebestová. Two-sided a posteriori error estimates for the DGMs for the heat equation. In *In A. Cangiani, R.L. Davidchack, E.H. Georgoulis, A. Gorban, J. Levesley, M. Tretyakov (eds.), Numerical Mathematics and Advanced Applications, Proceedings of the ENUMATH 2011 Conference, Leicester*. Springer, 2012.

[26] Shuyu Sun and Mary F. Wheeler. A posteriori error estimation and dynamic adaptivity for symmetric discontinuous Galerkin approximations of reactive transport problems. *Comput. Methods Appl. Mech. Engrg.*, 195(7-8):632–652, 2006.

[27] R. Verfürth. A posteriori error estimators for convection-diffusion equations. *Numer. Math.*, 80(4):641–663, 1998.

[28] R. Verfürth. Robust a posteriori error estimates for nonstationary convection-diffusion equations. *SIAM J. Numer. Anal.*, 43(4):1783–1802 (electronic), 2005.

[29] R. Verfürth. Robust a posteriori error estimates for stationary convection-diffusion equations. *SIAM J. Numer. Anal.*, 43(4):1766–1782 (electronic), 2005.

[30] Rüdiger Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Wiley-Teubner, Chichester-Stuttgart, 1996.